Dark Secrets of RF Design

Prof. Tom Lee

Stanford University

Director, DARPA Microsystems Technology Office

Inaugural IEEE SSCS Webinar

Why RF design is hard

- Can’t ignore parasitics.
- Can’t squander device power gain.
- Can’t tolerate much noise or nonlinearity.
- Can’t expect accurate models, but you still have to ship anyway.

Traditional RF design flow

- Don pointy wizard hat.
- Obtain chicken.
- Design first-pass circuit.
- Mumble obscure Latin incantations ("*semper ubi sub ubi*...*omnia pizza in octo partes divisa est...e pluribus nihil*").
- Test circuit; weep uncontrollably.
- Adjust chicken.

Dark secrets: A partial list

- MOSFETs: What your textbook didn't tell you
- The two-port noise model: Why care?
  - Optimum noise figure vs. maximum gain
- To match or not to match – that is the question
- Linearity and time-invariance revisited
- Mixers: Myths and noise
- Strange impedance behaviors (SIBs)

MOSFETs: What Your Textbooks May Not Have Told You

The standard lie

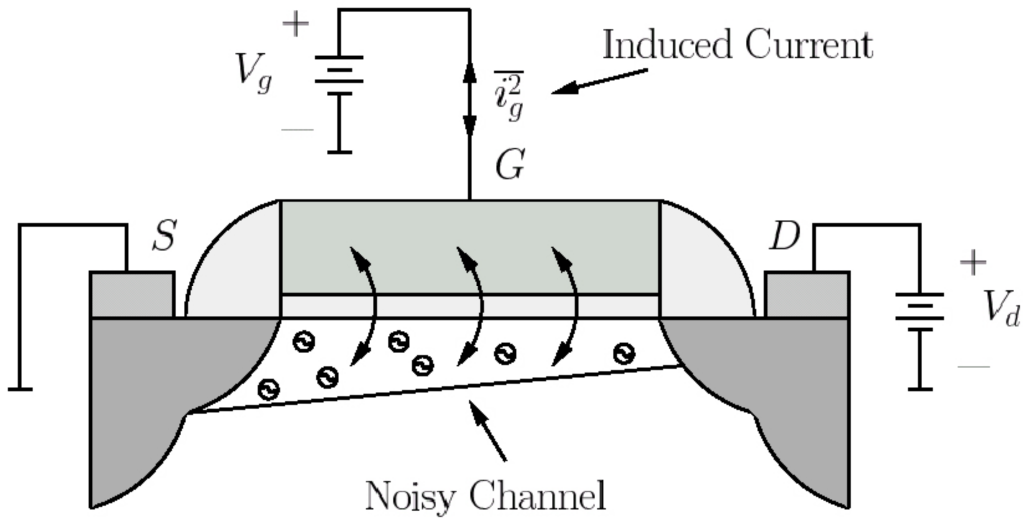
- "Gate-source impedance is a capacitor."
- Because zero power is thus needed to drive it, *any output at all, at any frequency, implies infinite power gain.* (The books usually omit that last part.)

The true story

- Gate-source impedance is *not* a pure capacitor.
- Phase shift associated with finite carrier transit speed means gate field does nonzero work on channel charge.
- Therefore, power gain is not infinite.
- *There is also noise associated with the dissipation.*

Noisy channel charge

- Fluctuations couple capacitively to both top and bottom gates.
  - Induces noisy gate currents. Bottom-gate term is ignored by most models and textbooks.



[Shaeffer]

- Gate noise current. $\overline{i_g^2} = 4kTB\delta\frac{(\omega C_{gs})^2}{5g_{d0}}$
- Real component of Y_g. $\mathrm{Re}\left[Y_g\right] = \frac{(\omega C_{gs})^2}{5g_{d0}}$

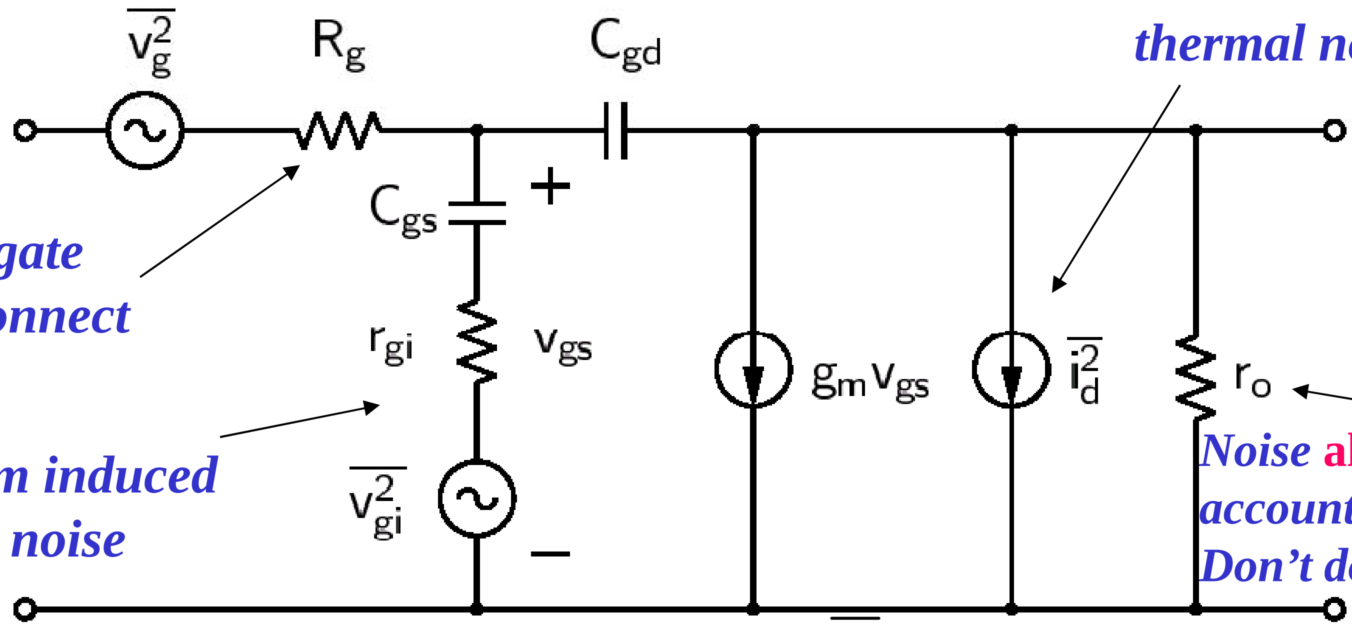
Sources of noise in MOSFETs

- (Thermally-agitated) channel charge.
  - Produces both drain *and* gate current noise.

- Interconnect resistance.
  - Series gate resistance R_g is very important.

- Substrate resistance.
  - Substrate thermal noise modulates back gate, augments drain current noise in some frequency range.

(All) FETs and gate noise: Basic model



$$\overline{i_d^2} = 4kT\gamma g_{d0}$$

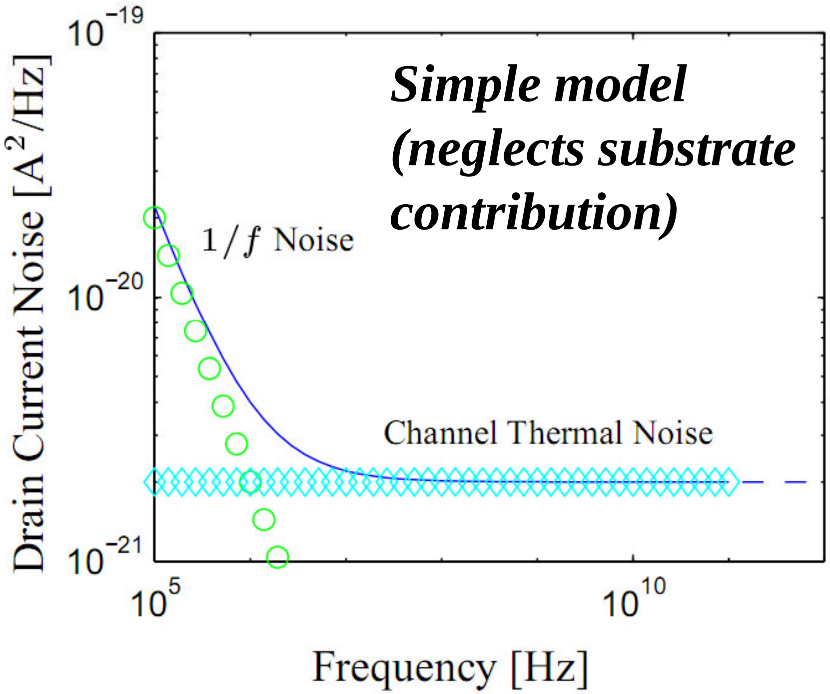
$$\overline{v_{gi}^2} = 4kTB\delta r_{gi}$$

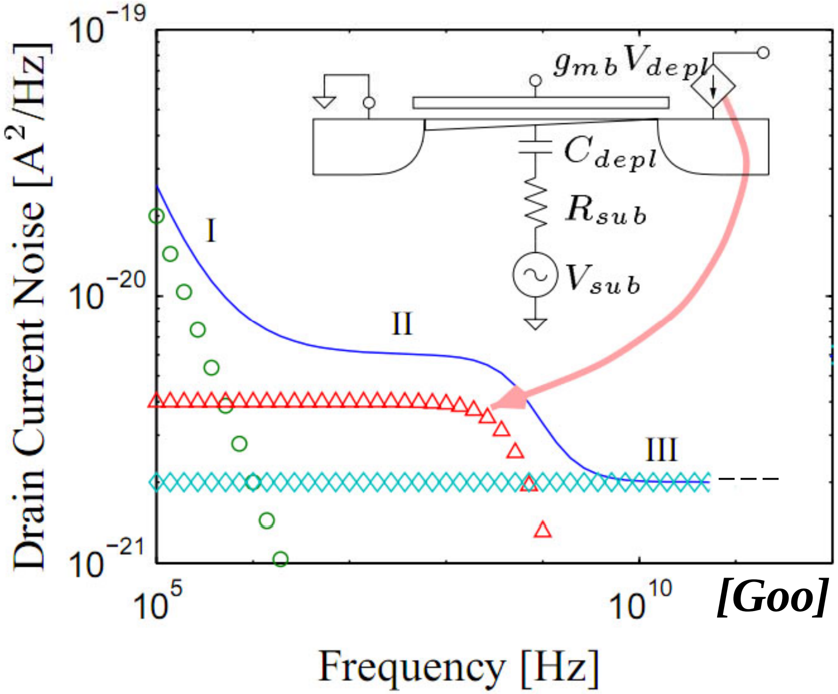
$$r_{gi} = \frac{1}{5g_{d0}}$$

(Note the placement of v_{gs}.)

Important: Common error is to define Vgs as across Cgs alone.

Substrate thermal noise





[Goo]

$$S_{i_d,sub} = \frac{4kTR_{sub}g_{mb}^2}{1+(\omega R_{sub}C_{depl})^2}$$

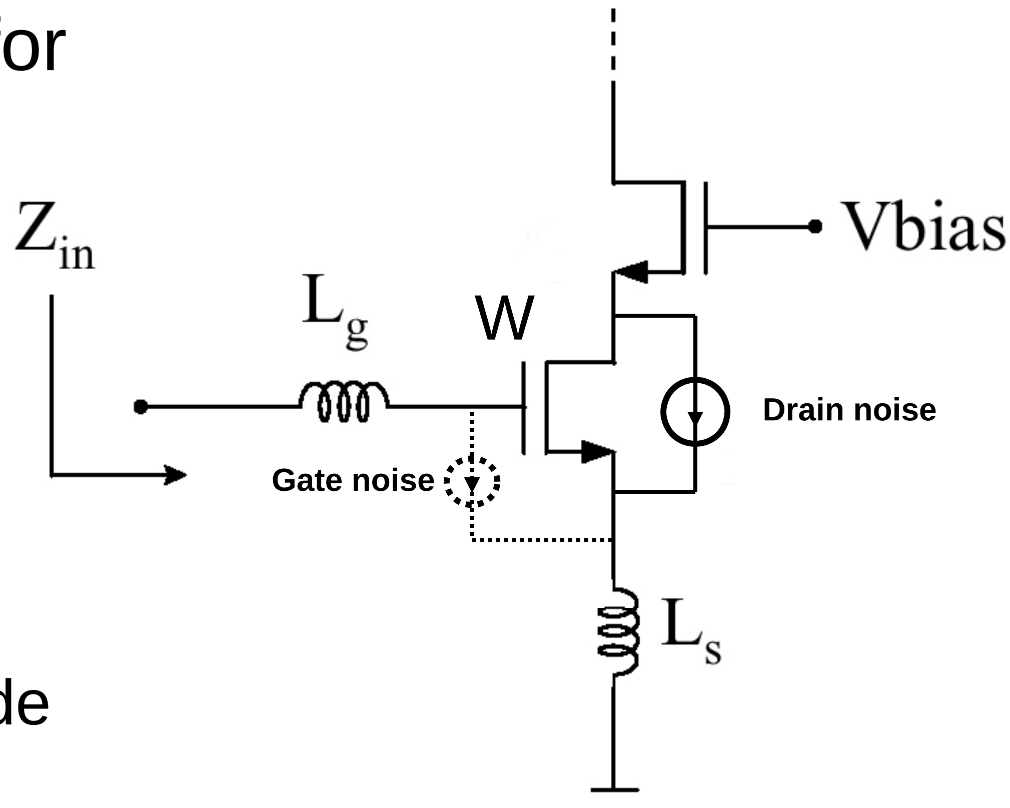
Substrate thermal noise controversy

- Measuring drain noise at different frequencies has led to confusion about the value of γ.
  - Measurements made below ~1GHz (i.e., in *Region II*) may reveal “excess” noise, and a sensitivity to the number of substrate taps, if wrong model is used.

- Early speculations that deep-submicron MOSFETs suffer from significant enhancement of γ not borne out.

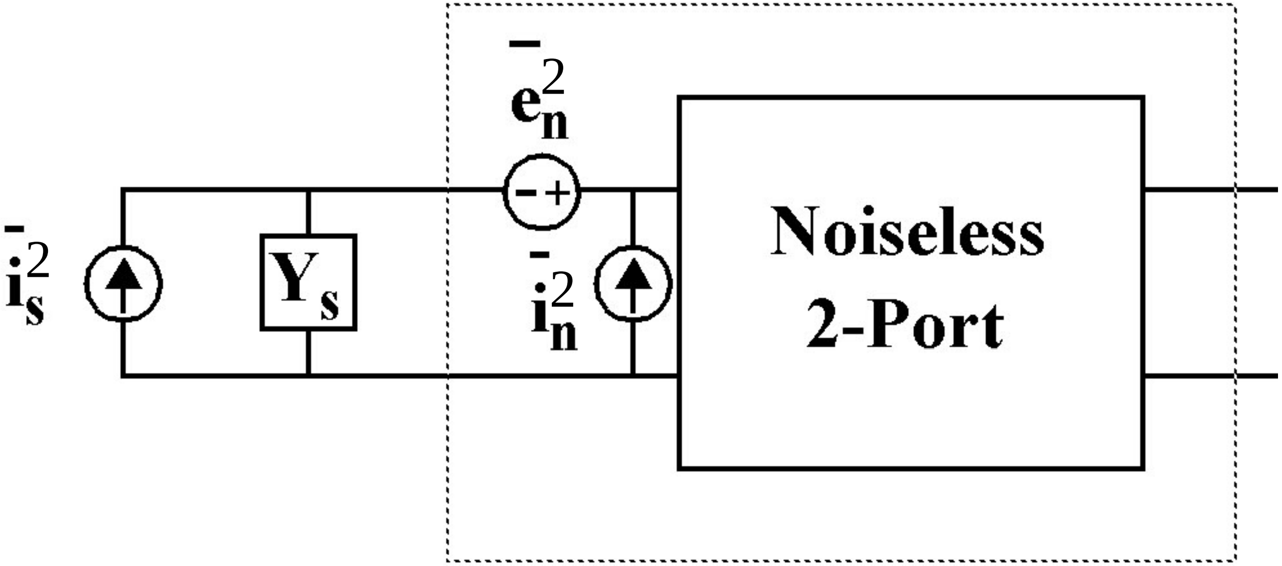
Gate noise is real; Murphy says so

- Let W → 0 while maintaining resonance and current density (for fixed f_T).
  - Gain stays fixed.
  - I_{bias} → 0.
- If you ignore gate noise:
  - Output noise → zero; absurd to consume zero power and provide noiseless gain.



The Two-Port Noise Model: Why Care?

Two-port noise model



$$F = \frac{\overline{i_s^2} + \overline{\left|i_n + Y_s e_n\right|^2}}{\overline{i_s^2}} = 1 + \frac{\overline{i_u^2} + \left|Y_c + Y_s\right|^2 \overline{e_n^2}}{\overline{i_s^2}}$$

- The *IRE* chose not to define *F* directly in terms of equivalent input noise sources. Instead:

Two-port noise model

Let $$R_n \equiv \frac{\overline{e_n^2}}{4kT\Delta f},$$

$$G_u \equiv \frac{\overline{i_u^2}}{4kT\Delta f},$$

and

$$G_s \equiv \frac{\overline{i_s^2}}{4kT\Delta f}.$$

Conditions for minimum noise figure

$$B_s = -B_c = B_{opt}$$

$$G_s = \sqrt{\frac{G_u}{R_n} + G_c^2} = G_{opt}$$

$$F_{min} = 1 + 2R_n[G_{opt} + G_c] = 1 + 2R_n\left[\sqrt{\frac{G_u}{R_n} + G_c^2} + G_c\right]$$

$$F = F_{min} + \frac{R_n}{G_s}\left[(G_s - G_{opt})^2 + (B_s - B_{opt})^2\right]$$

Important observation

- Minimum NF and maximum power gain occur for the same source Z *only if three miracles occur together*:
  - $G_c = 0$ (noise current has no component in phase with noise voltage); *and*
  - $G_u = G_n$ (conductance representing uncorrelated current noise equals the fictitious conductance that produces noise voltage); *and*
  - $B_c = B_{in}$. (correlation susceptance happens to be the same as the actual input susceptance)

To Match or Not to Match -- That is the Question

Impedance matching: Why?

- Conjugate match maximizes power transfer.

- Terminating a T-line in its characteristic impedance makes the input impedance length-independent.
  - Also minimizes peak voltage and current along line.

- Selecting and maintaining a standard impedance value (e.g., 50Ω) facilitates fixturing and instrumentation.

Impedance matching: Why not?

- Amplifiers generally exhibit best noise figure with a mismatch.
- Many amplifiers are more stable or robust (in the PVT sense) when mismatched.
- If power gain is not in short supply (and stability and noise are not a problem), may not need to match impedances, resulting in a simpler circuit.

Linearity and Time-Invariance: So What?

LTI, LTV and all that

- A system is linear if superposition holds.
- A system is TI if an input timing shift only shifts the timing of the output the same amount.
  - Shapes stay constant.
- If a system is LTI, it can only scale and phase-shift Fourier components.
  - Output and input frequencies are the same.
- If a system is LTV, input and output frequencies can be different, ***despite being linear***.
- If a system is nonlinear, input and output frequencies will generally differ.

Mixers are supposed to be linear!

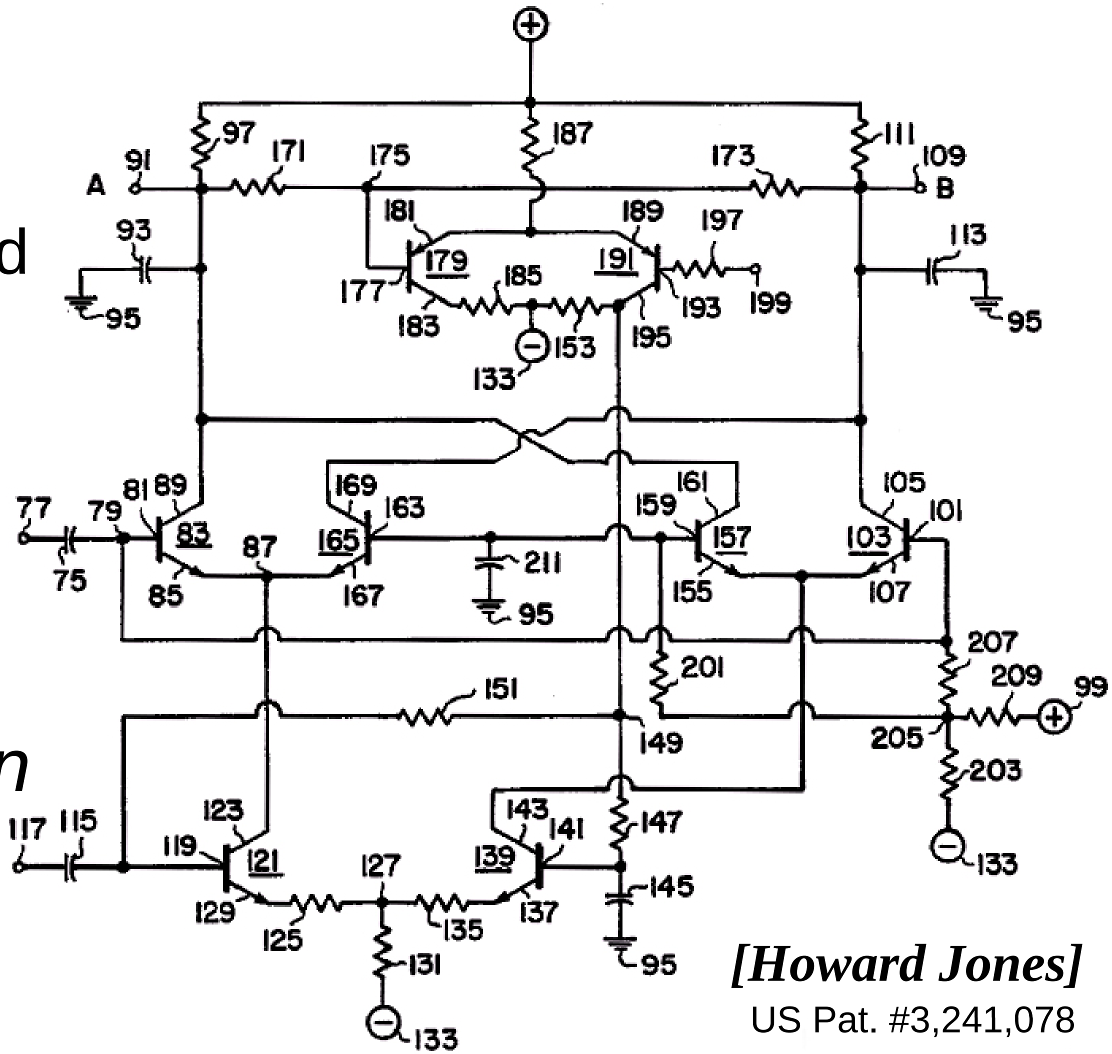
- But they are *time-varying* blocks.
  - Ignore textbooks and papers that say "mixers are nonlinear…" Mixers are nonlinear in the same way amplifiers are nonlinear: *Undesirably*.

- Significantly noisier than LNAs for reasons that will be explained shortly. NF values of 10-15dB are not unusual.

- Main function of an LNA is usually to provide enough gain to overcome mixer noise.

First: This is *not* a Gilbert mixer

- This is a *Jones* mixer.
  - Most textbooks and papers (still) wrongly call this a Gilbert cell.
- A true Gilbert cell is a *current-domain* circuit, and uses predistortion for linearity.



[Howard Jones]
US Pat. #3,241,078

The mixer: An LTV element

- Whether Gilbert, Jones or Smith, modern mixers depend on *commutation* of currents or voltages.

- We idealize mixing as the equivalent of multiplying the RF signal by a square-wave LO.
  - Single-balanced mixer: RF signal is unipolar.
  - Double-balanced mixer: RF signal is DC-free.

- Mixing is ideally *linear*: Doubling the input (RF) voltage should double the output (IF) voltage.

A multiplier is an ideal mixer

- Key relationship is:

$$A\cos\omega_1 t\cos\omega_2 t = \frac{A}{2}[\cos(\omega_1+\omega_2)t + \cos(\omega_1-\omega_2)t]$$

- Can be thought of as an amplifier with a time-varying amplification factor.

Mixer noise figure

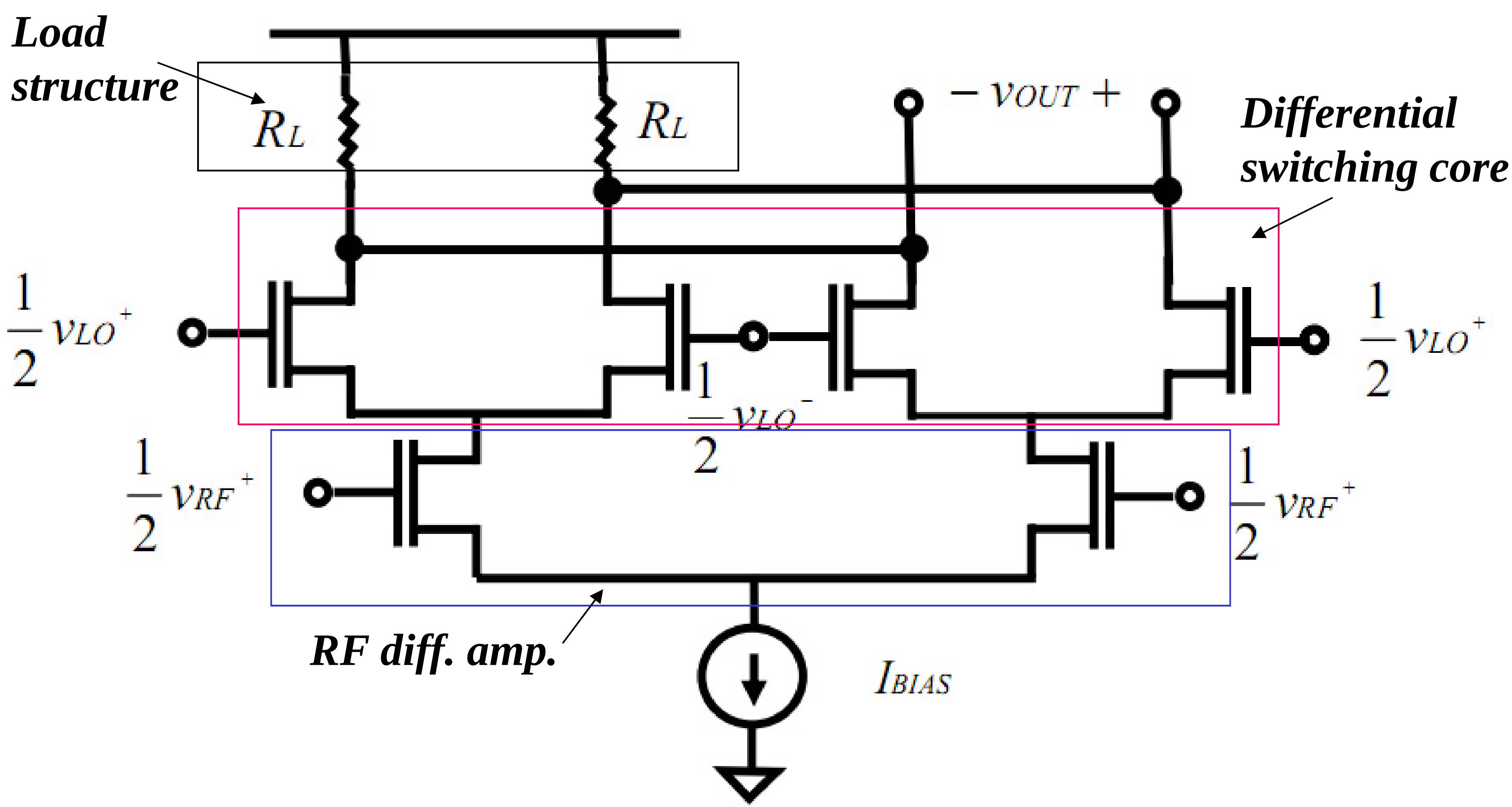
- Noise figure of mixers is worse than for LNAs for several reasons.
  - Noise originating from different RF bands can translate to the same IF.
  - Transconductor is usually optimized more for linearity than for noise.
  - Switching core contributes significant noise in practical mixers.

Mixer noise figure: DSB v. SSB

- Because noise from two different sidebands (desired RF and its image, located $2f_{IF}$ away) can convert to the same IF, need to be careful about defining NF.

- If both sidebands contain signal (and noise), we report DSB NF. If signal is present in only one sideband, we report SSB NF.
  - If noise gains are constant, DSB NF = SSB NF – 3dB.
  - Because DSB NF is lower, it gets reported more frequently. *Beware*.

Sources of noise in mixers



Mixer noise

- Load structure is at the output, so its noise adds to the output directly; it undergoes no frequency translations.
  - If 1/*f* noise is a concern, use PMOS transistors or poly resistor loads.
- Transconductor noise appears at same port as input RF signal, so it translates in frequency the same way as the RF input.

Dark secret: Switching noise can *dominate*

- Instantaneous switching not possible.
  - Noise from switching core can actually *dominate*.
  - Common-mode capacitance at tail nodes of core reduces effectiveness of large LO amplitudes.

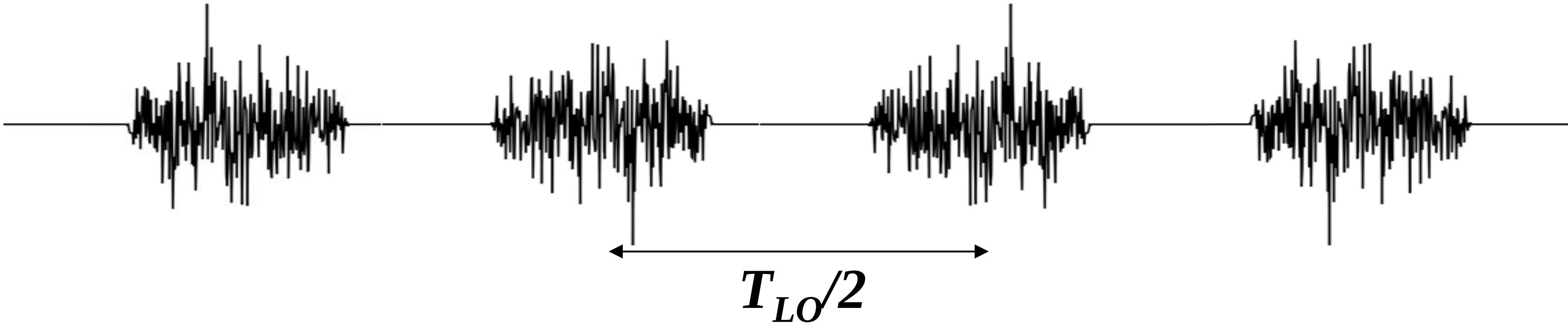
- Periodic switching of core is equivalent to sampling core noise at (twice) the LO rate.
  - Frequency translations occur due to this self-mixing.

Noise contribution of switching core

- As switching transistors are driven through the switching instant, they act as a differential pair for a brief window of time t_s.
  - During this interval, the switching transistors transfer their drain noise to the output.
  - Changing drain current implies a changing PSD for the noise; it is cyclostationary.

Noise contribution of switching core

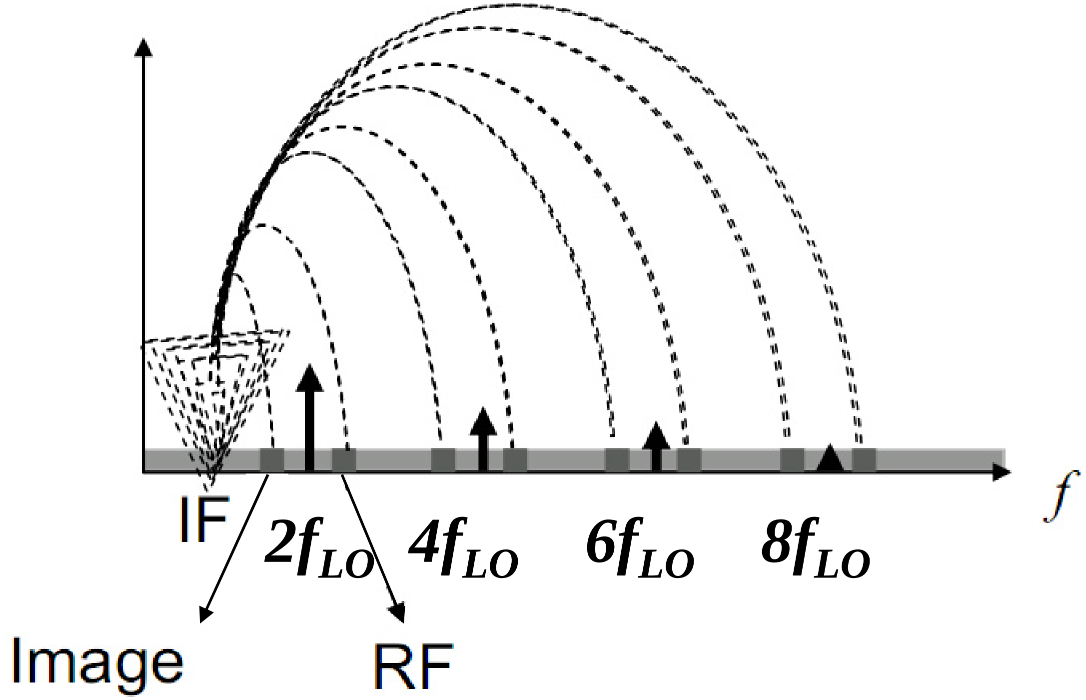
- The noise contributed by the switching core appears as follows:



- Mathematically equivalent to multiplying a stationary noisy waveform by a sampling pulse train with fundamental frequency $2f_{LO}$.

Noise contribution of switching core

- Noise at $2nf_{LO}$ +/- f_{IF} will therefore translate to the IF. This noise folding helps explain the relatively poor noise figure of mixers.



Terrovitis mixer noise figure equation

- A simplified analytical approximation for the SSB noise figure of a Jones mixer is

important

$$F_{SSB} \approx \frac{\alpha}{c^2} + \frac{2\gamma g_m \alpha + 4\gamma \overline{G} + G_L}{c^2 g_m^2 R_S}$$

- Here, g_m is the transconductance of the bottom differential pair; G_L is the conductance of the load; R_S is the source resistance, and γ is the familiar drain noise parameter.
  - See *[Terrovitis]* for more complete version.

Terrovitis mixer noise figure equation

- The parameter $\overline{G}$ is the time-averaged transconductance of each pair of switching transistors. For a plain-vanilla Jones mixer,

$$\overline{G} \approx \frac{2I_{BIAS}}{\pi V_{LO}}$$

- The parameter α is related to the sampling aperture, and has an approximate value

$$\alpha \approx 1 - \frac{4}{3} t_s f_{LO}$$

Terrovitis mixer noise figure equation

- The parameter *c* is directly related to the effective aperture, and is given by

$$c \approx \frac{2}{\pi}\left[\frac{\sin(\pi t_s f_{LO})}{\pi t_s f_{LO}}\right]$$

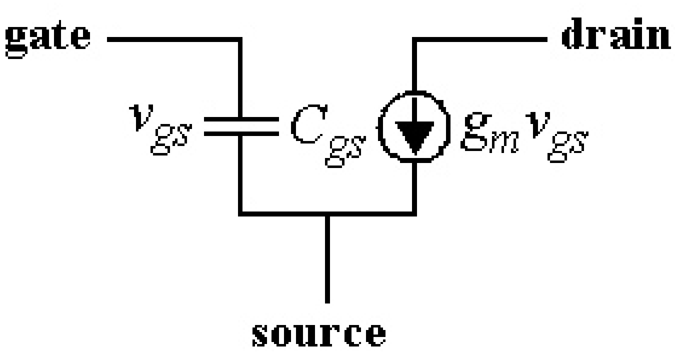
- This parameter asymptotically approaches $2/\pi$ in the limit of infinitely fast switching.

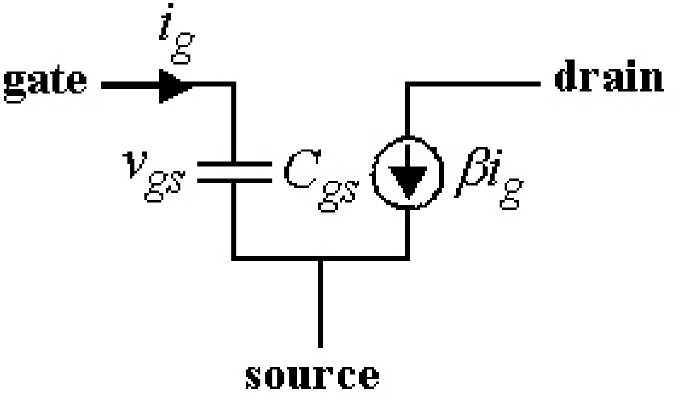
When Good Amplifiers Go Bad:

Strange Impedance Behaviors

First: Some simple transistor models

- Can use *either* gate-source voltage *or* gate current as independent control variable



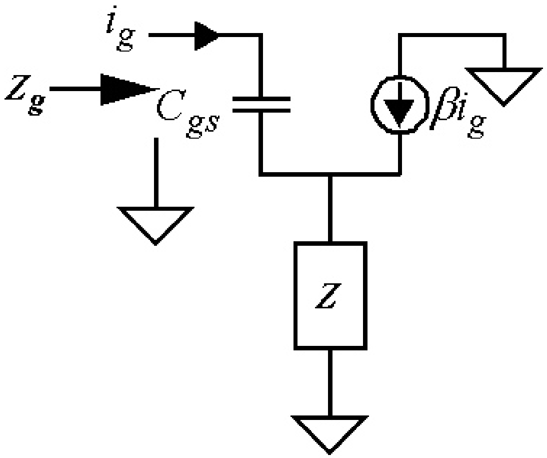


- Models are fully equivalent as long as we choose

$$\beta = \frac{g_m v_{gs}}{i_g} = \frac{g_m}{sC_{gs}} = \frac{\omega_T}{j\omega} = -j\frac{\omega_T}{\omega}$$

View from the gate

- Consider input impedance of the following at $\omega \ll \omega_T$:



$$Z_g = \frac{1}{j\omega C_{gs}} + Z(\beta + 1) \approx \frac{1}{j\omega C_{gs}} + Z(-j\frac{\omega_T}{\omega})$$

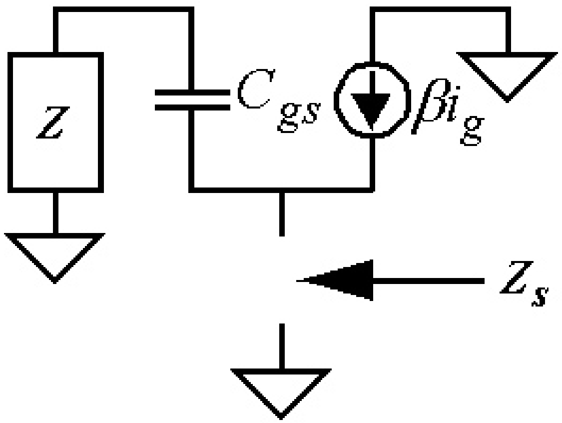
- The non-intuitive behavior comes from the second term: The impedance *Z* gets multiplied by a (negative) imaginary constant.

What does multiplication by $-j\omega_T/\omega$ do?

- Turns R into capacitance of value $1/\omega_T R$.
- Turns L into resistance of value $\omega_T L$.
- Turns C into *negative* resistance of value $-\omega_T/\omega^2 C$.

View from the source

- Now consider input impedance of the following:



$$Z_s = \frac{\frac{1}{j\omega C_{gs}} + Z}{\beta + 1} \approx \frac{1}{g_m} + Z\left(j\frac{\omega}{\omega_T}\right)$$

- This time, *Z* gets multiplied by a +*j* factor.

What does multiplication by $+j\omega/\omega_T$ do?

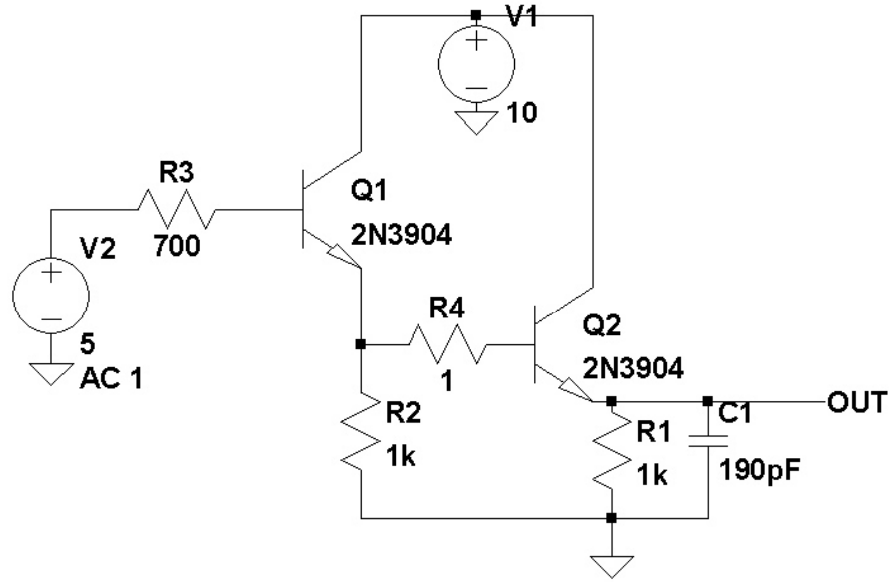
- Turns R into inductance of value R/ω_T.
- Turns C into resistance of value $1/\omega_T C$.
- Turns L into negative resistance of value $-\omega^2 L/\omega_T$.

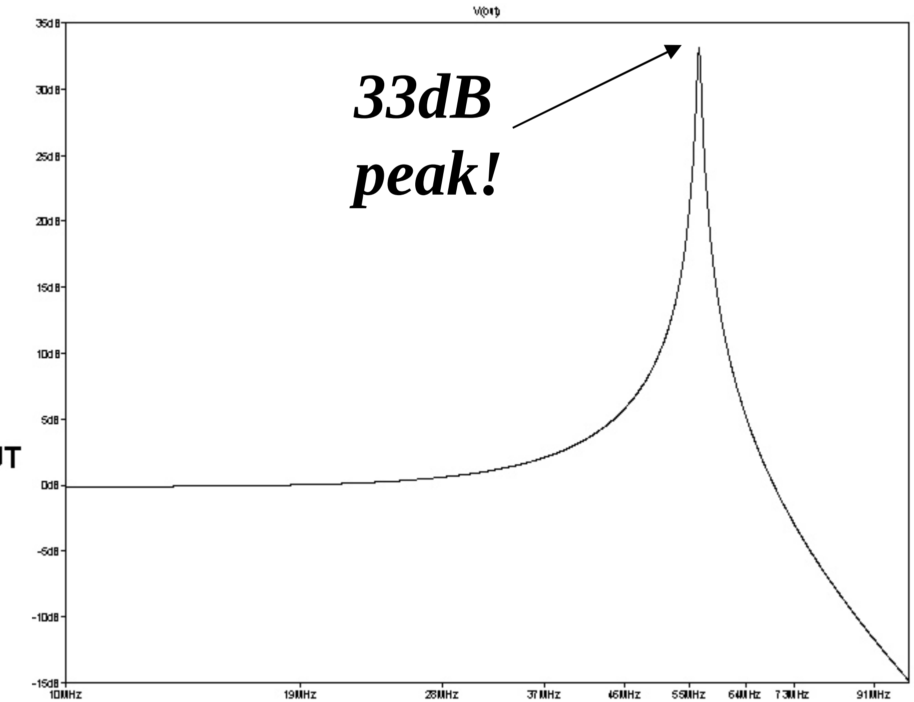
Why SIBs are strange

- Apparent weirdness arises because the current gain is imaginary.
- Quadrature phase shift associated with imaginary current gain causes impedances to change *character*, not just magnitude.
- The strangeness evaporates once you spend a little time studying where it comes from.

SIBs example: Follower cascade

- Familiar circuit has surprising and terrifying behavior:





Summary

- RF circuits are certainly complex, but that shouldn't make us concede defeat.

- Everything is explicable; it's not magic!

- So throw away the pointy hat, free the chickens, quit babbling in Latin, and stop weeping uncontrollably.

References

[Goo] J.S. Goo, *High Frequency Noise in CMOS Low-Noise Amplifiers*, Doctoral Dissertation, Stanford University, August 2001.

[Jones] H. E. Jones, US Pat. #3,241,078, "Dual Output Synchronous Detector Utilizing Transistorized Differential Amplifiers," issued March 1966.

[Lee] *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd edition, Cambridge U. Press, 2004.

[Shaeffer] D. Shaeffer and T. Lee, "A 1.5-V, 1.5-GHz CMOS Low *Noise* Amplifier," *IEEE J. Solid-State Circuits*, v.32, pp. 745-758, 1997.

[Terrovitis] M. T. Terrovitis and R. G. Meyer, "Noise in Current-Commutating CMOS Mixers," *IEEE Journal of Solid-State Circuits,* vol. 34, No. 6, June 1999.